Ultrasensitive N-photon interferometric autocorrelator

Zili Zhou*,[1] Giulia Frucci,[1] Francesco Mattioli,[2] Alessandro Gaggero,[2] Roberto
Leoni,[2] Saeedeh Jahanmirinejad,[1] Thang Ba Hoang,[1] and Andrea Fiore[1]

[1]*COBRA Research Institute, Eindhoven University of Technology,
P.O. Box 513, 5600 MB Eindhoven, The Netherlands*
[2]*Istituto di Fotonica e Nanotecnologie (IFN), CNR, via Cineto Romano 42, 00156 Rome, Italy*

We demonstrate a novel method to measure the Nth-order (N=1, 2, 3, 4) interferometric auto-
correlation with high sensitivity and temporal resolution. It is based on the combination of linear
absorption and nonlinear detection in a superconducting nanodetector, providing much higher ef-
ficiency than methods based on all-optical nonlinearities. Its temporal resolution is only limited
by the quasi-particle energy relaxation time, which is directly measured to be in the 20 ps range
for the NbN films used in this work. We present a general model of interferometric autocorrela-
tion with these nonlinear detectors and discuss the comparison with other approaches and possible
improvements.

The temporal correlation functions of various orders are of fundamental importance in the classical and quantum description of optical fields. The first-order (field) autocorrelation function describes temporal coherence and therefore spectral linewidth, the second-order (intensity) autocorrelation is used to measure the temporal properties of pulsed sources and to distinguish quantum and classical fields, while the measurement of higher-order autocorrelation is more sensitive to coherence features (e.g. photon bunching) of the light field [1] and can be used to determine the asymmetry of light pulses [2]. While the first-order autocorrelation function is easily measured using an interferometer and a linear detector [3], the measurement of higher-order correlation functions requires a process which is nonlinear in the intensity $I(t)$. In interferometric autocorrelators, the normalized second-order correlation function $g^{(2)}(\tau) = \langle I(t)I(t+\tau)\rangle/\langle I(t)\rangle^2$ is usually measured by using either second-harmonic generation (SHG) in a nonlinear crystal, followed by a linear detector [4], or two-photon absorption (TPA) in the detector itself [5, 6]. In both cases, the detector measures the square of the total intensity at the output of the interferometer $\langle I_{\mathrm{tot}}^2(t)\rangle \propto \langle I^2(t)\rangle + 2\langle I(t)I(t+\tau_{\mathrm{d}})\rangle$ where τ_{d} is the delay between the two arms of the interferometer, together with interference terms which are sensitive to the phase properties of the beam. While these approaches offer very high temporal resolution, since the related processes are nearly instantaneous, their sensitivity is limited by the low nonlinear susceptibilities involved in the SHG or TPA process. Due to the even lower relevance of higher-order optical nonlinearities, they cannot be effectively applied to the measurement of the autocorrelations of order $N > 2$. An alternative approach consists of combining linear optical detection with nonlinear processing

in the electrical read-out, e.g. in a correlation card, as in the Hanbury Brown and Twiss interferometer [7]. In this case, sensitive single-photon detectors can be used, however the temporal resolution is limited ≥ 100 ps by the jitter of the detector output and of the amplification and correlation electronics [8]. In addition, differently from interferometric autocorrelation, these approaches do not provide any information on the phase properties.

In this work we report a novel approach to the measurement of the interferometric autocorrelation of order $N \geq 2$, which is based on the combination of *linear absorption* and *nonlinear detection* in a single device. The general principle consists of absorbing incident photons in a linear absorber (i.e. a material where the absorption probability per unit time is proportional to light intensity), which produces an output pulse only if two or more photons are absorbed within a certain time interval in the fs or ps range. We show the implementation of this concept in a superconducting nanodetector, where the nonlinearity is widely tunable by varying the bias current. We directly measure the temporal dynamics of the nonlinearity in the ps range and we attribute it to the relaxation dynamics of photo-created quasi-particles (QP). We show its application to the measurement of up to fourth-order interferometric autocorrelation, observing an extremely high sensitivity related to the linear absorption process and to the low detector noise.

The superconducting nanodetector consists of a nanoscale constriction in a superconducting wire (see Fig.1). The device used in this work is based on a 4.3 nm-thick NbN film (critical temperature T_{c}=10.2 K) sputtered on GaAs substrate and has a constriction size of about 150 nm, patterned by electron beam lithography and reactive-ion etching. The nanodetector is biased with a current I_{b} smaller than the superconducting critical current I_{c}. Similarly to nanowire superconducting single-photon detectors [10], the absorption of one or more photons produces a non-equilibrium population of QP in the nanodetector's active region, locally suppressing the superconductivity and increasing the probability

of vortex crossing, which can result in the transition to the normal state [11, 12]. The detection probability is a strong function of the absorbed energy (i.e. to the number of photons) and of the bias current I_b, so that the nanodetector can be set to respond to $\geq N$ photons by choosing I_b [13]. A full tomographic characterization of the multi-photon response of the nanodetector under illumination with short pulses has been reported in Ref. [14].
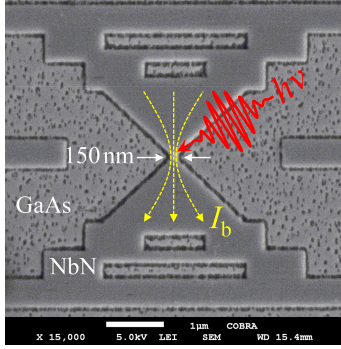

FIG. 1: Scanning electron microscope image of a NbN superconducting nanodetector.

In this study we focus on the temporal characteristics of the nanodetector's multi-photon response and on its application as an interferometric autocorrelator. As we focus here on the application to the characterization of classical light sources, we describe the detection process as a sequence of single-photon absorption events using the semi-classical photo-detection theory. Assuming an incident light pulse with cycle-average intensity $I_{in}(t)$ on the nanodetector, the probability of creating a hot-spot within a time interval $[t + dt]$ is equal to $\zeta \cdot S \cdot I_{in}(t) \cdot dt$, where ζ is related to the absorptance η_{abs} by $\zeta = \eta_{abs}/h\nu$ and S is the active area. In the two-photon regime, for example, the click probability P_{click} (assumed $\ll 1$) is equal to the probability that two photons are absorbed, weighted by a function $\eta(\tau_{12})$ of the time difference τ_{12} between absorption of the first and the second photon at times t_1 and t_2 respectively,

$$
\begin{aligned}
P_{click} &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \eta(t_2 - t_1)\zeta S I_{in}(t_1)\zeta S I_{in}(t_2) \cdot dt_1 \cdot dt_2 \\
&= \zeta^2 S^2 \int_{-\infty}^{+\infty} \eta(\tau_{12}) \int_{-\infty}^{+\infty} I_{in}(t)I_{in}(t+\tau_{12}) \cdot dt \cdot d\tau_{12}, \quad (1)
\end{aligned}
$$

The $\eta(\tau_{12})$, named nonlinear response function (NRF) in the following, depends on the QP dynamics in the superconductor and determines the intrinsic response time τ_{ND} of the nanodetector. The $\eta(\tau_{12})$ is expected to decay from a maximum value of $\eta(0)$(which depends on I_b) to 0 for $\tau_{12} \gg \tau_{ND}$. In general, the detection probability in

the N-photon regime ($N \geq 2$) is,

$$
\begin{aligned}
P_{click} = \zeta^N S^N \int_{-\infty}^{+\infty} &\eta_N(\tau_{1N}, \tau_{2N}, \cdots, \tau_{N-1,N}) \int_{-\infty}^{+\infty} I_{in}(t) \\
&I_{in}(t+\tau_{1N})I_{in}(t+\tau_{2N})\cdots I_{in}(t+\tau_{N-1,N}) \\
&\cdot dt \cdot d\tau_{1N} \cdot d\tau_{2N} \cdots d\tau_{N-1,N}
\end{aligned} \quad (2)
$$

Where $\tau_{iN} = t_N - t_i$ denotes the difference between the absorption times of the ith- and the last photon.

When the incident pulse width is much larger than τ_{ND}, Eq.(2) is approximated as $P_{click} \propto \int_{-\infty}^{+\infty} I_{in}^N(t) \cdot dt$. When a nanodetector is placed at the output of a Michelson interferometer, the input intensity reads $I_{in}(t) \propto |E(t) + E(t+\tau_d)|^2$ and the P_{click} in N-photon regime is proportional to the Nth-order interferometric autocorrelations given by $P_{click}(\tau_d) \propto \int_{-\infty}^{+\infty} [|E(t) + E(t+\tau_d)|^2]^N \cdot dt$. In the two-photon regime, by filtering out the interference terms we obtain $P_{click}(\tau_d) \propto \langle I^2(t) \rangle + 2\langle I(t)I(t+\tau_d) \rangle$, and therefore the $g^{(2)}(\tau_d)$, similarly to the SHG- and TPA-based autocorrelators.

The NRF can be measured by probing the autocorrelator with short pulses. For pulse duration much shorter than τ_{ND}, the response of the nanodetector placed at the output of a Michelson interferometer can be found from Eq.(2). In the two-photon regime for example, after filtering out the interference terms, Eq.(1) becomes $P_{click}(\tau_d) \propto \eta(0) + \eta(\tau_d) + f(\tau_d)$ [15]. The function $f(\tau_d)$ depends on the degree of first-order coherence of the input light and is different from zero only for delays shorter than the coherence time. For longer delays, the $P_{click}(\tau_d)$, normalized by its value at $\tau_d \gg \tau_{ND}$, is expected to vary as $P_{click}(\tau_d)/P_{click}(\infty) = 1 + \eta(\tau_d)/\eta(0)$, so that the NRF and therefore the autocorrelator's timing resolution can be measured.

The NRF was first measured by sending 1.6 ps pulses from an optical parametric oscillator (OPO) at $\lambda = 1.13\mu m$ into a fiber-based Michelson interferometer and then to a nanodetector held at a temperature of 1.2 K using a lensed fiber producing a spot with an e^{-2} diameter of 5 μm. The delay τ_d in the interferometer is controlled by a motorized delay line (coarse control) and a fiber stretcher (fine control). The I_c of the device was about 26 μA. During the measurement, the nanodetector was set in different photon regimes by choosing a proper I_b. As shown in Fig. 2(a), the count rate (CR) was measured as a function of the light power at different I_b. The solid lines with slopes of 1.04, 2.06, 3.06 and 3.99 are the fits to the measured data in log-log scale, in the power ranges where the one-, two-, three- and four-photon response was dominant [13]. We chose one point at each of the four photon regimes, and measured the CR as a function of τ_d, normalized by its values at long delays, as shown in Fig. 2(b). The data points near $\tau_d = 0$, where the measured autocorrelation is sensitive to the first-order coherence, are not shown in the

plot and were not considered in the fit since they introduce additional fitting error [15]. At $I_\mathrm{b} = 18.0$ μA, the CR is independent of the delay, apart from the short-delay interference fringes, as expected, since the detector is working in the linear regime. When the I_b was lowered to 12.5 μA, 9.7 μA and 8.4 μA (corresponding to the two-, three- and four-photon regime, respectively), a maximum is observed at zero delay. Since the width of these peaks is much wider than the OPO pulse width, the measurement probes the intrinsic response of the detector. In particular, the two-photon normalized response in Fig. 2(b) directly provides the $1 + \eta(\tau_\mathrm{d})/\eta(0)$ dependence. The $\eta(\tau)$ is determined by the thermalization and relaxation processes of the photo-created QP and by the functional dependence of the P_click on the QP concentration [11, 12]. Indeed, the QP population first grows as the electron population thermalizes via electron-electron scattering, in a timescale of few ps, then decays due to electron-photon interaction and phonon escaping to the substrate [16]. This QP decay is expected to determine the τ_ND in the N-photon regime, since the I_b is chosen so that P_click is high only when all QP produced by the N photons are present at the same time. As P_click is a strong function of the QP number, involving many microscopic parameters, a fit of the measured $\eta(\tau)$ using a microscopic model would not be reliable. Instead, we introduce an empirical Gaussian NRF defined as $\eta(\tau_{12}) = \eta(0)\exp[-(\tau_{12}/\tau_\mathrm{ND})^2]$. The fits to ten measured two-photon autocorrelation traces provide a τ_ND value of 20.4 ± 0.8 ps. In order to fit the $N > 2$ traces we further assume that the multi-photon response factorizes as $\eta_N(\tau_{1N}, \tau_{2N}, \cdots, \tau_{N-1,N}) = \eta(\tau_{1N})\eta(\tau_{2N})\cdots\eta(\tau_{N-1,N})$. This is reasonable if one assumes that η_N has an approximately exponential dependence on the total QP concentration after absorption of the Nth-photon, as suggested by the vortex-assisted photo-detection model [11, 12], and that the QP relaxation time does not depend on QP concentration. Using the τ_ND value extracted from the two-photon autocorrelation traces as described above, $P_\mathrm{click}(\tau_\mathrm{d})$ was calculated for the three- and four-photon regimes from Eq.(2) without additional fitting parameters and shows excellent agreement with the experiment (Fig. 2(b)), which provides a strong experimental support to our model. The value of τ_ND measured here at 1.2K is comparable but shorter than the hot-electron energy relaxation time in NbN microbridge measured by electro-optic sampling [17] and by Terahertz spectroscopy [18] - this difference may be related to the difference in film thickness and film properties.

With the knowledge of the temporal resolution, autocorrelation experiments have been performed on pulses generated by a gain-switched 1.3μm diode laser with 10MHz repetition rate and about 70ps pulse width. The one-, two-, three- and four-photon regimes were first found by choosing different I_b. At each I_b, CR were recorded as a function of τ_d in the N-photon (N=1, 2,
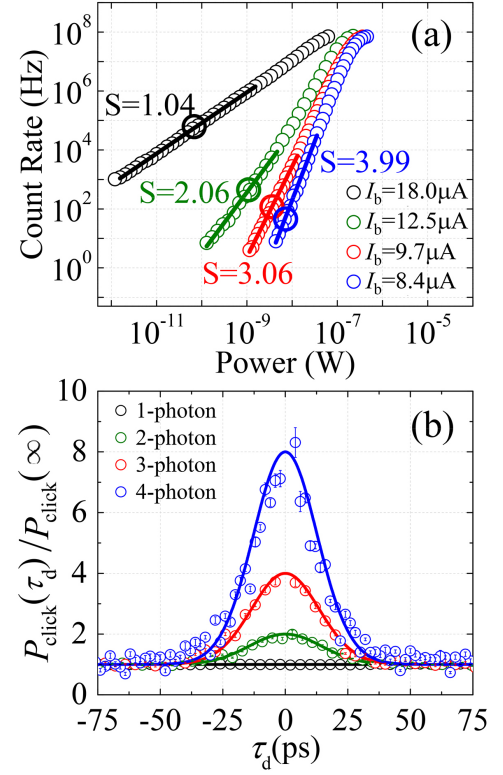

FIG. 2: (a) CR as a function of the light power for different I_b. The fitting lines with slopes of 1.04, 2.06, 3.06 and 3.99 indicate the one-, two-, three- and four-photon regimes respectively. (b) Normalized CR (circles) as a function of τ_d, measured when the detector was working at the four chosen points (marked by large circles in (a)). The solid lines are calculation results (see text).

3, 4) regime as shown in Fig. 3(a-d), respectively. The fringe contrast ratio is observed to increase with N, in good agreement with the theoretical values of 2, 8, 32 and 128 for N=1, 2, 3 and 4, respectively. In the one-photon regime (Fig. 3(a)), the normalized first-order autocorrelation function $g^{(1)}(\tau_\mathrm{d})$ was calculated from the visibility of the interference fringes [3]. For $N > 1$, higher-order intensity autocorrelation traces were obtained by applying a low-pass filter to the interferograms in Fig. 3(b-d). In order to explain the experimental results, we assumed that the incident light was a Gaussian pulse with a linear chirp. The electric field is modeled as $E(t) = E_0\exp[-4\ln(2)(1 + iA)t^2/\tau_\mathrm{p}^2]$, where E_0 is the field amplitude, A is the linear chirp parameter [19] and τ_p is the full width at half maximum (FWHM) of the pulse. By fitting the measured $|g^{(1)}(\tau_\mathrm{d})|$ in Fig. 3(a) together with the second-order intensity autocorrelation (low-pass trace) in Fig. 3(b), A and τ_p were determined to be 5.3 and 70.6 ps. The third and fourth-order interferometric autocorrelations were then calculated based on Eq.(2) without additional fitting parameters. Their low-pass traces show a good agreement with the experiments, considering the very simplified assumption for the
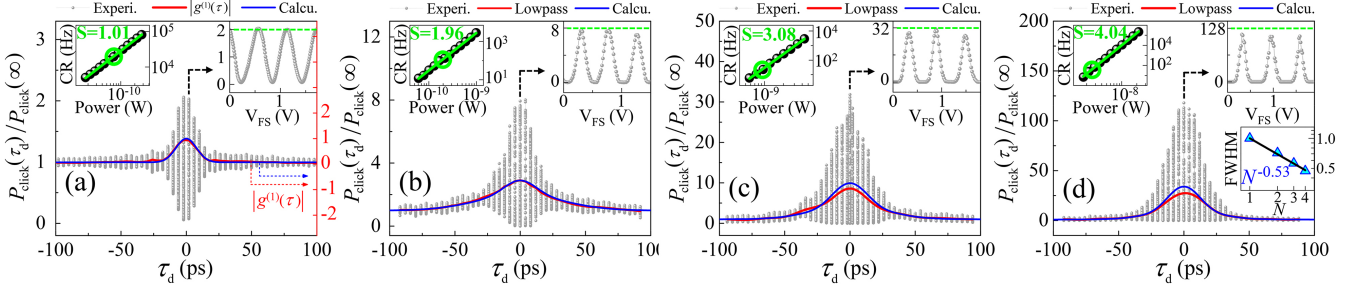
FIG. 3: (a-d) Normalized CR as a function of τ_d in the different regimes at I_b of 18.0 μA, 14.0μA, 11.5μA and 9.6μA respectively. The $|g^{(1)}(\tau_d)|$ (red line in (a)) and low-pass curves (red lines in (b-d)) agree well with the calculations (blue lines). Upper-left insets: CR as a function of incident power in log-log scale. The fitting slopes of 1.01, 1.96, 3.08 and 4.04 indicate the one-, two-, three- and four-photon regimes, respectively. The green circles indicate the points chosen for the measurements in the main panel. Upper-right insets: Expanded view of the interference fringes at small delays, plotted as a function of the voltage V_{FS} applied to the fiber stretcher. Lower-right inset in (d): FWHM of the fringes normalized by their period extracted from the upper-right insets in (a-d) as a function of N, showing a $1/\sqrt{N}$ dependence.

chirp. An enlarged view of each interferogram is shown as upper-right inset in each panel of Fig. 3, showing a clear narrowing of the fringes for increasing N. As shown in the lower-right inset of Fig. 3(d), the FWHM of the fringes, normalized by their period, scales as approximately $1/\sqrt{N}$, which is a characteristic of multi-photon interferometry [20] and further confirms our conclusions.

A comparison of the sensitivity of our nanodetector to the detection schemes used in conventional autocorrelators provides useful insight. For input pulses longer than τ_{ND}, Eq.(2) can be written as $P_{click} = C_{ND} \int_{-\infty}^{+\infty} P_{in}^2(t) dt$, where $P_{in} = I_{in}S$ is the incident power and $C_{ND} = \sqrt{\pi}\eta_{abs}^2\eta(0)\tau_{ND}/(h\nu)^2$ represents a nonlinear response efficiency. A similar expression for P_{click} can be derived for nonlinear detectors based on SHG or TPA, where the corresponding C_{SHG} and C_{TPA} values are related to the SHG nonlinear conversion efficiency and to the TPA coefficient β_2, respectively. In all three cases, the sensitivity can be defined by imposing that the CR is equal to the dark count rate R_{dark} in the detector. For a periodic sequence of square pulses with repetition rate f_{rep} the condition $f_{rep}P_{click} \geq R_{dark}$ leads to $P_{pk}P_{av} \geq R_{dark}/C$, where P_{pk} and P_{av} are the peak and average power, respectively. The two-photon autocorrelation trace in Fig. 3(b) was taken at $P_{pk}P_{av} = 5.6 \times 10^{-17}$W^2, about seven orders of magnitude lower than the minimum reported $P_{pk}P_{av}$ for TPA using a GaAs PM tube [21] and about two orders of magnitude lower than the lowest $P_{pk}P_{av}$ reported in the second-order interferometric correlations using SHG [22]. Using the measured values of $\eta_{abs} = 1.5 \times 10^{-4}$, $\eta(0) \approx 0.5$ [14] and $R_{dark} = 1$Hz, we derive a sensitivity of $P_{pk}^{min}P_{av}^{min} \approx 5.8 \times 10^{-20}$W^2, corresponding to $\sim$4 photons per pulse in our experiment. In higher-photon regime the advantage of using linear absorption is even larger. Indeed, the three-photon autocorrelation shown in Fig. 3(c) was performed at an average power of about 1nW, corresponding to $P_{pk}^2 P_{av} \approx 2.0 \times 10^{-21}$W^3, an improvement of about twenty-one orders of magnitude over

Ref. [23]. N-photon interferometric autocorrelation for $N > 3$ has not been reported before to the best of our knowledge.

The very high nonlinear response of the nanodetector, can be directly traced to the finite size and time duration of the hot-spot created by the *real* absorption of one photon, as compared to virtual transitions involved in TPA. Indeed, a two-photon detection is triggered if the second photon is absorbed within the volume and time duration of the hot-spot created by the first photon. This shows that a compromise exists between C_{ND} and τ_{ND}: for the nanodetector the τ_{ND} is determined by the QP relaxation time, while in TPA it is related to the lifetime of the virtual states associated to the TPA transition, of the order of fs [6]. A similar compromise exists in SHG-based autocorrelators, where higher conversion efficiency requires a longer SHG crystal translating into a smaller phase-matching bandwidth and lower temporal resolution [24, 25]. Interestingly, in an absorption thickness of 4.3 nm the nanodetector reaches a nonlinear efficiency $C \approx 1.7 \times 10^{19}$W^{-1}J^{-1}, several orders of magnitude higher than μm-thick TPA absorbers and comparable to cm-long SHG crystals, showing the giant effective nonlinearity achieved through the combination of linear absorption and nonlinear detection. We note that a key advantage as compared to SHG-based autocorrelators is the wide optical bandwidth, limited only by the requirement to operate in the desired N-photon regime, which can be easily adjusted by varying the I_b. This in principle enables the measurement of interferometric autocorrelation from the visible to the mid-infrared.

The sensitivity of our autocorrelator is presently limited by the low η_{abs}, related to the spatial mismatch between the incoming beam and the nanodetector's active area and to the small thickness of the NbN film. By focusing the beam with a high-numerical aperture lens, a much higher absorptance $\eta_{abs} \approx 10^{-2}$ can be achieved [26]. The integration of a plasmonic antenna and of a bot-

tom reflector [27] could increase η_{abs} to the 10^{-1} range, leading to $P_{\mathrm{pk}}^{\mathrm{min}} P_{\mathrm{av}}^{\mathrm{min}} \approx 10^{-25} \mathrm{W}^2$ for the two-photon autocorrelation. On the other hand, increasing the detector area (as done in meander nanowire detectors [28]) results in $\eta_{\mathrm{abs}} \propto L$ and $\eta(0) \propto l_{\mathrm{hs}}/L$ (L is the nanowire length, l_{hs} is the hot-spot length) so that C_{ND} scales linearly with η_{abs}. The nanoscale nature of the detector is therefore crucial to reaching the ultimate sensitivity. Finally, we note that the 20ps temporal resolution in our experiments is limited to the QP relaxation time in NbN films and could be much improved using different superconducting materials, such as high-T_{c} Y-Ba-Cu-O films, where relaxation times ~ 1 ps were observed [29], opening the way to the characterization of Nth-order correlation functions in few-ps range with unprecedented sensitivity.

---

[1] M. Aßmann *et al.*, Science **325**, 297 (2009).
[2] P. Langlois and E. P. Ippen, Opt. Lett. **24**, 1868 (1999).
[3] M. Fox, *Quantum optics: an introduction* (Oxford University Press: New York, 2006), Chapter 2.
[4] H. P. Weber, J. Appl. Phys. **38**, 2231 (1967).
[5] B. R. Mollow, Phys. Rev. **175**, 1555 (1968).
[6] F. Boitier *et al.*, Nat. Photon. **5**, 267 (2009).
[7] R. Hanbury Brown and R. Q. Twiss, Nature (London) **177**, 27 (1956); **178**, 1046 (1956).
[8] The fast correlation technique demonstrated in Ref. [1, 9] also belongs to this category, but suffers from the low sensitivity of streak cameras in the near-infrared.
[9] J. Wiersig *et al.*, Nature (London) **460**, 245 (2009).
[10] G. N. Gol'tsman *et al.*, Appl. Phys. Lett. **79**, 705 (2001).
[11] A. D. Semenov *et al.*, Physica C **468**, 627 (2008).
[12] L. N. Bulaevskii *et al.*, Phys. Rev. B **85**, 014505 (2012).
[13] D. Bitauld *et al.*, Nano Lett. **10**, 2977 (2010).
[14] J. J. Renema *et al.*, Opt. Express **20**, 2806 (2012).
[15] Supplemental Material can be found in the following pages
[16] A. D. Semenov *et al.*, Physica C **351**, 349 (2001).
[17] K. S. Il'in *et al.*, Appl. Phys. Lett. **76**, 2752 (2000).
[18] M. Beck *et al.*, Phys. Rev. Lett. **107**, 177007 (2011).
[19] J. C. M. Diels *et al.*, Appl. Opt. **24**, 1270 (1985).
[20] E. Yablonovitch and R. B. Vrijen, Opt. Eng. **38**, 334 (1999).
[21] J. M. Roth *et al.*, Opt. Lett. **27**, 2076 (2002).
[22] C. S. Hsu *et al.*, Opt. Lett. **36**, 2611 (2011).
[23] Y. Wei *et al.*, Opt. Lett. **36**, 2372 (2011).
[24] A. M. Weiner, IEEE J. Quantum Electron. **19**, 1276 (1983).
[25] S. D. Yang *et al.*, Appl. Opt. **46**, 6759 (2007).
[26] This value is obtained by assuming the diameter of the Gaussian beam spot decreases from 5μm to 0.5μm by using high-numerical aperture lens, and the nanodetector's active area is 150nm×150nm.
[27] A. Gaggero *et al.*, Appl. Phys. Lett. **97**, 151108 (2010).
[28] A. Verevkin *et al.*, Appl. Phys. Lett. **80**, 4687 (2002).
[29] M. Lindgren *et al.*, IEEE J. Sel. Top. Quantum Electronics **2**, 668 (1996).

Supplemental Material:

In the general case, the response of the nanodetector,

which is set in the two-photon regime and placed at the output of the interferometer, is found using Eq. (1) and reads,

$$P_{\text{click}}(\tau_{\text{d}}) \propto 2\int\limits_{-\infty}^{+\infty} \eta(\tau_{12})G^{(2)}(\tau_{12}) \cdot \mathrm{d}\tau_{12} + \int\limits_{-\infty}^{+\infty} \eta(\tau_{12})G^{(2)}(\tau_{12}+\tau_{\text{d}}) \cdot \mathrm{d}\tau_{12} + \int\limits_{-\infty}^{+\infty} \eta(\tau_{12})G^{(2)}(\tau_{12}-\tau_{\text{d}}) \cdot \mathrm{d}\tau_{12}$$

$$+2\mathrm{Re}\left[\int\limits_{-\infty}^{+\infty} \eta(\tau_{12})\int\limits_{-\infty}^{+\infty} E^*(t)E(t+\tau_{12})E(t+\tau_{\text{d}})E^*(t+\tau_{12}+\tau_{\text{d}}) \cdot \mathrm{d}t \cdot \mathrm{d}\tau_{12}\right]$$

$$+\int\limits_{-\infty}^{+\infty} \eta(\tau_{12})F_1(\tau_{12},\omega\tau_{\text{d}}) \cdot \mathrm{d}\tau_{12} + \int\limits_{-\infty}^{+\infty} \eta(\tau_{12})F_2(\tau_{12},2\omega\tau_{\text{d}}) \cdot \mathrm{d}\tau_{12} \qquad (S1)$$

The first term on the right side of Eq. (S1) is a constant with respect to τ_{d}. The second and the third terms represent the convolution of the NRF with the second-order correlation function defined as $G^{(2)}(\tau) = \int_{-\infty}^{+\infty} I(t)I(t+\tau) \cdot \mathrm{d}t$, while the fourth term is sensitive to the phase-coherence of the input beam and gives rise to an additional peak in the filtered autocorrelation signal

at small delays. The last two terms on the right side of Eq. (S1) are interference terms which vary with $\omega\tau_{\text{d}}$ or $2\omega\tau_{\text{d}}$. Since they average to zero they can be easily filtered out using a low-pass filter.

In the case where the pulse is much shorter than τ_{ND}, Eq. (S1) is written as,

$$P_{\text{click}}(\tau_{\text{d}}) \propto \eta(0) + \eta(\tau_{\text{d}}) + \frac{\eta(0)\left|G^{(1)}(\tau_{\text{d}})\right|^2}{\int\limits_{-\infty}^{+\infty} G^{(2)}(\tau_{12}) \cdot \mathrm{d}\tau_{12}} + \text{interference terms} \qquad (S2)$$

Where $G^{(1)}(\tau_{\text{d}})$ is the first-order correlation function defined as $G^{(1)}(\tau_{\text{d}}) = \int_{-\infty}^{+\infty} E^*(t)E(t+\tau_{\text{d}}) \cdot \mathrm{d}t$. The third term on the right side of Eq. (S2), which comes from the fourth term in Eq. (S1), depends on the first-order coherence properties of the input beam, but, differently from the interference terms, does not disappear upon filtering and produces a peak in the filtered autocorrelation at zero delay.

In order to confirm this prediction, we studied the short delay window in the experiment where the OPO short pulses (1.6ps, 1.13μm) were used to characterize the NRF of the nanodetector in the two photon regime. To this aim, the delay τ_{d} was controlled by a motorized delay line (coarse control) and a fiber stretcher (FS) (fine control). A driving voltage V_{FS} (zigzag wave, ±5V, 20mHz) was applied on the fiber stretcher to change the optical path difference between the two arms. At each coarse delay, the counter recorded the count rate as a function of the V_{FS}. Due to the short pulse width of OPO, the interference fringes were observed at small coarse delays

and an additional peak appeared in a small window near zero delay (blue points as shown in Fig. 4(a), also shown in an enlarged view in the inset), on top of the broad peak related to the detector NRF, as observed in Fig. 2(b) of the main text and replotted as green circles in Fig. 4(a). A theoretical comparison based on Eq. (S1) is shown in Fig. 4(b). It considers a 20.4ps intrinsic response time of the nanodetector, and a probe pulse of 1.6ps duration and 1.13μm wavelength as used in the experiments, showing a good agreement with the measurements. We note that this narrow peak appears also in the filtered autocorrelation, shown as red lines in the insets of Fig. 4(a) and (b). However, as indicated by Eq. (S2), its width relates to the first-order correlation function, and therefore to the coherence time, and not to the pulse width. Indeed, we checked numerically that in a chirped pulse the narrow peak width scales with the chirp parameter. This confirms that τ_{ND} represents the limit to the temporal resolution for the measurement of the second-order correlation function.
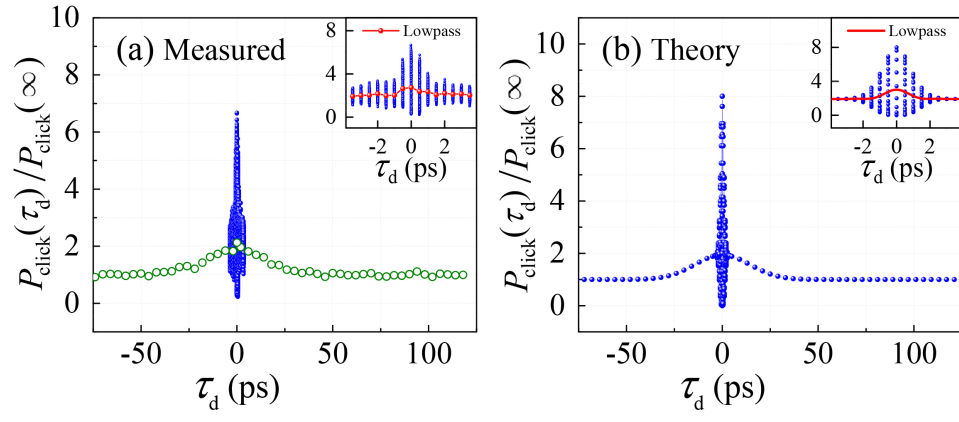
FIG. 4: (a) The measured count rate as a function of the time delay, normalized by its value at long delays. (b) Simulation based on Eq. (S1). An enlarged view at small delays is shown in the insets.